

(19) 日本国特許庁 (J P)

(12) 特 許 公 報 (B 2)

(11) 特許番号

特許第3360030号

(P3360030)

(45) 発行日 平成14年12月24日 (2002. 12. 24)

(24) 登録日 平成14年10月11日 (2002. 10. 11)

(51) Int.Cl. <sup>7</sup>	識別記号	F I
G 0 6 K 9/03		G 0 6 K 9/03 C
9/62	6 4 0	9/62 6 4 0 A
	6 5 0	6 5 0 A
9/72		9/72 B

請求項の数12(全 16 頁)

(21) 出願番号	特願平10-222913	(73) 特許権者	000004226 日本電信電話株式会社 東京都千代田区大手町二丁目3番1号
(22) 出願日	平成10年8月6日 (1998. 8. 6)	(72) 発明者	荒木 麻衣 東京都新宿区西新宿三丁目19番2号 日 本電信電話株式会社内
(65) 公開番号	特開2000-57257(P2000-57257A)	(72) 発明者	宮本 信夫 東京都新宿区西新宿三丁目19番2号 日 本電信電話株式会社内
(43) 公開日	平成12年2月25日 (2000. 2. 25)	(72) 発明者	鈴木 章 東京都新宿区西新宿三丁目19番2号 日 本電信電話株式会社内
審査請求日	平成12年12月22日 (2000. 12. 22)	(74) 代理人	100087848 弁理士 小笠原 吉義 (外1名)
		審査官	廣瀬 文雄

最終頁に続く

(54) 【発明の名称】 文字認識装置および文字認識方法および文字認識方法をプログラムの形で記録した記録媒体

1

(57) 【特許請求の範囲】

【請求項1】 名称文字コード列と数字文字コード列の対が格納されている単語テーブルと、  
 名称文字と数字文字とを含む画像を入力する画像入力部と、  
 上記画像から全ての文字列画像を切り出す文字列切り出し部と、  
 上記文字列画像の各々から全ての文字画像を切り出す文字切り出し部と、  
 上記文字列画像に含まれる全ての文字画像の認識を行い、認識結果文字コード列を出力する文字認識部と、  
 上記文字列切り出し部で得られた全文字列について、上記認識結果文字コード列と、上記単語テーブル内の全名称文字コード列との間の認識結果を出力する名称単語照合部と、

2

上記文字列切り出し部で得られた全文字列について、上記認識結果文字コード列と、上記単語テーブル内の全数字文字コード列との間の認識結果を出力する数字単語照合部と、  
 上記文字列切り出し部で得られた全文字列について、上記文字切り出し部より出力された文字画像の大きさのばらつきをもとに、名称文字列と見なしたときの認識結果信頼度を算出する名称文字列信頼度計算部と、  
 上記文字列切り出し部で得られた全文字列について、上記認識結果文字コード列中の数字の割合をもとに、数字文字列と見なしたときの認識結果信頼度を算出する数字文字列信頼度計算部と、  
 上記文字列切り出し部で得られた全文字列について、上記名称文字列信頼度計算部により算出された認識結果信頼度と、上記数字文字列信頼度計算部により算出された

10

認識結果信頼度とから、名称文字列として見たときの重み係数と、数字文字列として見たときの重み係数とを算出する重み計算部と、

名称文字列に対する認識結果と数字文字列に対する認識結果とを、重みを加えて統合する重み付け統合部とを有することを特徴とする文字認識装置。

【請求項2】 上記文字列切り出し部により得られた文字列の全ての対について、上記名称単語照合部により得られた一致文字数に、上記重み計算部により算出された名称文字列として見たときの重み係数を乗じた値と、上記数字単語照合部により得られた一致文字数に、上記重み計算部により算出された数字文字列として見たときの重み係数を乗じた値との和をスコアとして算出し、最大スコアを与える名称文字コード列を名称候補として出力する重み付け統合部とを有することを特徴とする請求項1記載の文字認識装置。

【請求項3】 上記文字列切り出し部で得られた全文字列の中で、大きさが最大の文字を含む文字列を名称文字列候補として抽出し、その文字列の認識結果文字コード列を名称文字列候補として出力する名称文字列候補抽出部と、

上記文字列切り出し部で得られた全文字列の中で、上記認識結果文字コード列中の数字の数が最も多いものを数字文字列候補として抽出し、その文字列の認識結果文字コード列を数字文字列候補として出力する数字文字列候補抽出部とを有することを特徴とする請求項1記載の文字認識装置。

【請求項4】 上記数字文字列候補抽出部より出力された上記数字文字列候補の上記認識結果文字コード列に対して、数字文字列の書式を基に、単語照合に使用するための数字部分のみを抽出する数字部分抽出部を有することを特徴とする請求項3記載の文字認識装置。

【請求項5】 名称文字コード列と数字文字コード列の対が格納されている単語テーブルを用い、名称文字と数字文字とを含む画像を入力する画像入力段階と、

上記画像から全ての文字列画像を切り出す文字列切り出し段階と、

上記文字列画像の各々から全ての文字画像を切り出す文字切り出し段階と、

上記文字列画像に含まれる全ての文字画像の認識を行い、認識結果文字コード列を出力する文字認識段階と、上記文字列切り出し段階で得られた全文字列について、上記認識結果文字コード列と、上記単語テーブル内の全名称文字コード列との間の認識結果を出力する名称単語照合段階と、

上記文字列切り出し段階で得られた全文字列について、上記認識結果文字コード列と、上記単語テーブル内の全数字文字コード列との間の認識結果を出力する数字単語照合段階と、

上記文字列切り出し段階で得られた全文字列について、上記文字切り出し段階より出力された文字画像の大きさのばらつきをもとに、名称文字列と見なしたときの認識結果信頼度を算出する名称文字列信頼度計算段階と、

上記文字列切り出し段階で得られた全文字列について、上記認識結果文字コード列中の数字の割合をもとに、数字文字列と見なしたときの認識結果信頼度を算出する数字文字列信頼度計算段階と、

上記文字列切り出し段階で得られた全文字列について、上記名称文字列信頼度計算段階により算出された認識結果信頼度と、上記数字文字列信頼度計算段階により算出された認識結果信頼度とから、名称文字列として見たときの重み係数と、数字文字列として見たときの重み係数とを算出する重み計算段階と、名称文字列に対する認識結果と数字文字列に対する認識結果とを、重みを加えて統合する重み付け統合段階とを有することを特徴とする文字認識方法。

【請求項6】 上記文字列切り出し段階により得られた文字列の全ての対について、上記名称単語照合段階により得られた一致文字数に、上記重み計算段階により算出された名称文字列として見たときの重み係数を乗じた値と、上記数字単語照合段階により得られた一致文字数に、上記重み計算段階により算出された数字文字列として見たときの重み係数を乗じた値との和をスコアとして算出し、最大スコアを与える名称文字コード列を名称候補として出力する重み付け統合段階とを有することを特徴とする請求項5記載の文字認識方法。

【請求項7】 上記文字列切り出し段階で得られた全文字列の中で、大きさが最大の文字を含む文字列を名称文字列候補として抽出し、その文字列の認識結果文字コード列を名称文字列候補として出力する名称文字列候補抽出段階と、

上記文字列切り出し段階で得られた全文字列の中で、上記認識結果文字コード列中の数字の数が最も多いものを数字文字列候補として抽出し、その文字列の認識結果文字コード列を数字文字列候補として出力する数字文字列候補抽出段階とを有することを特徴とする請求項5記載の文字認識方法。

【請求項8】 上記数字文字列候補抽出段階より出力された上記数字文字列候補の上記認識結果文字コード列に対して、数字文字列の書式を基に、単語照合に使用するための数字部分のみを抽出する数字部分抽出段階を有することを特徴とする請求項7記載の文字認識方法。

【請求項9】 名称文字コード列と数字文字コード列の対が格納されている単語テーブルを用い、名称文字と数字文字とを含む画像を入力する画像入力段階と、

上記画像から全ての文字列画像を切り出す文字列切り出し段階と、

上記文字列画像の各々から全ての文字画像を切り出す文

字切り出し段階と、

上記文字列画像に含まれる全ての文字画像の認識を行い、認識結果文字コード列を出力する文字認識段階と、上記文字列切り出し段階で得られた全文字列について、上記認識結果文字コード列と、上記単語テーブル内の全名称文字コード列との間の認識結果を出力する名称単語照合段階と、

上記文字列切り出し段階で得られた全文字列について、上記認識結果文字コード列と、上記単語テーブル内の全数字文字コード列との間の認識結果を出力する数字単語照合段階と、

上記文字列切り出し段階で得られた全文字列について、上記文字切り出し段階より出力された文字画像の大きさのばらつきをもとに、名称文字列と見なしたときの認識結果信頼度を算出する名称文字列信頼度計算段階と、

上記文字列切り出し段階で得られた全文字列について、上記認識結果文字コード列中の数字の割合をもとに、数字文字列と見なしたときの認識結果信頼度を算出する数字文字列信頼度計算段階と、

上記文字列切り出し段階で得られた全文字列について、上記名称文字列信頼度計算段階により算出された認識結果信頼度と、上記数字文字列信頼度計算段階により算出された認識結果信頼度とから、名称文字列として見たときの重み係数と、数字文字列として見たときの重み係数とを算出する重み計算段階と、

名称文字列に対する認識結果と数字文字列に対する認識結果とを、重みを加えて統合する重み付け統合段階とを有する文字認識方法をプログラムの形で記録したことを特徴とする記録媒体。

【請求項10】 上記文字列切り出し段階により得られた文字列の全ての対について、上記名称単語照合段階により得られた一致文字数に、上記重み計算段階により算出された名称文字列として見たときの重み係数を乗じた値と、上記数字単語照合段階により得られた一致文字数に、上記重み計算段階により算出された数字文字列として見たときの重み係数を乗じた値との和をスコアとして算出し、最大スコアを与える名称文字コード列を名称候補として出力する重み付け統合段階とを有する文字認識方法をプログラムの形で記録したことを特徴とする請求項9記載の記録媒体。

【請求項11】 上記文字列切り出し段階で得られた全文字列の中で、大きさが最大の文字を含む文字列を名称文字列候補として抽出し、その文字列の認識結果文字コード列を名称文字列候補として出力する名称文字列候補抽出段階と、

上記文字列切り出し段階で得られた全文字列の中で、上記認識結果文字コード列中の数字の数が最も多いものを数字文字列候補として抽出し、その文字列の認識結果文字コード列を数字文字列候補として出力する数字文字列候補抽出段階とを有する文字認識方法をプログラムの形

で記録したことを特徴とする請求項9記載の記録媒体。

【請求項12】 上記数字文字列候補抽出段階より出力された上記数字文字列候補の上記認識結果文字コード列に対して、数字文字列の書式を基に、単語照合に使用するための数字部分のみを抽出する数字部分抽出段階を有する文字認識方法をプログラムの形で記録したことを特徴とする請求項11記載の記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、街中や駅の看板に書かれている文字を認識し、看板広告主を特定するようにした、文字認識装置および文字認識方法および文字認識方法をプログラムの形で記録した記録媒体に関するものである。

【0002】

【従来の技術】従来の文字認識に当っては、認識対象の文字列画像から文字を1文字ずつ切出して認識するものであった。しかしながら、看板のように、文字が多種多様な装飾文字で構成されている場合には、個々の文字の認識精度は十分ではなかった。そこで、認識対象文字列として出現しやすい文字列の単語テーブルを予め用意しておき、認識結果文字列との照合の結果、一致文字数の多い文字列があれば、それを認識結果として出力する単語照合処理がしばしば用いられてきた。

【0003】

【発明が解決しようとする課題】単語照合処理の効果を上げるためには、出現の可能性のある文字列を出来るだけ多く単語テーブルに登録する必要がある。しかし、その一方で登録文字列数を増やすほど類似した文字列が登録され、単語照合処理では認識結果を確定できないという問題があった。例えば、入力文字列画像として「A信託銀行」が入力されたとき、たとえ「A信託銀行」が表に登録されていても、他に「B信託銀行」が登録されていると、先頭文字「A」自体が誤認識されたときには、それが「A信託銀行」なのか「B信託銀行」なのか決定することができなかった。

【0004】本発明は上記事情に鑑みてなされたもので、その目的とするところは、看板における広告主などを認識するに当って高精度な認識を可能にすることにある。

【0005】

【課題を解決するための手段】上記の課題を達成するために、本発明は看板などから得られる複数文字列の文字認識結果と電話番号などの文字認識結果とを用いて、認識結果の信頼度を反映した照合結果の統合を図ることにより、精度良く看板の広告などを認識する文字認識装置を提供する。

【0006】即ち看板などの中に現れる複数の文字列を、広告主の名称を表す文字列、または電話番号を表す文字列として抽出し、それぞれの文字列で、あらかじめ

システムが有する単語テーブルに登録されている広告主文字列、電話番号文字列とのマッチングを行う。広告主の認識結果と電話番号の認識結果のどちらにより重みをおいて判断するかを、抽出した文字列から計算した信頼度によって看板画像ごとに動的に変動すること、より高い信頼度で広告主候補を確定する。

【0007】なお、以下看板の例を用いて説明する。より詳細にはそのための手段として、広告主名文字コード列と電話番号文字コード列との対が格納されている単語テーブルと、広告主名と電話番号とを含む看板画像を入力する画像入力部と、入力された看板画像から全ての分断のない一行を文字列として切り出す文字列切り出し部と、全文字列について、一文字ずつの文字画像領域を求めて文字画像を切り出す文字切り出し部と、それぞれの文字列に含まれる全ての文字画像の認識を行って認識結果文字コード列を出力する文字認識部とを持つ。更に上記そのための手段としての、文字列切り出し部で得られた全文字列について、認識結果文字コード列と単語テーブル内の全広告主名文字コード列との間の一致文字数を計数して出力する広告主単語照合部と、文字列切り出し部で得られた全文字列について、認識結果文字コード列と単語テーブル内の全電話番号文字コード列との間の一致文字数を計数して出力する電話番号単語照合部と、文字列切り出し部より出力された文字画像の大きさのばらつきをもとに、広告主文字列と見なしたときの認識結果信頼度を算出する広告主文字列信頼度計算部と、文字列切り出し部で得られた全文字列について、認識結果文字コード列中の数字の割合をもとに、電話番号と見なしたときの認識結果信頼度を算出する電話番号文字列信頼度計算部と、文字列切り出し部で得られた全文字列について、広告主文字列信頼度計算部により算出された認識結果信頼度と、電話番号文字列信頼度計算部により算出された認識結果信頼度とから、広告主文字列として見たときの重み係数と電話番号として見たときの重み係数とを算出する重み計算部とを持つ。そして更に上記そのための手段としての、文字列切り出し部により得られた文字列の全ての対について、広告主単語照合部により得られた一致文字数に、重み計算部により算出された広告主文字列として見たときの重み係数を乗じた値と、電話番号単語照合部により得られた一致文字数に、重み計算部により算出された電話番号文字列として見たときの重み係数を乗じた値との和をスコアとして算出し、最大スコアを与える広告主名文字コード列を広告主候補として出力する重み付け統合部とより構成される。

【0008】広告主の単語照合結果と電話番号の単語照合結果の一致文字数に重み付けをし、和を取ることで単語照合結果の統合を行うため、それぞれの文字列から十分な文字認識結果が得られなかった場合も、両方の照合結果で補完し合うことにより、高い広告主特定精度が得

られる。また、認識結果の信頼度でそれぞれの文字列の照合結果に重み付けをすることによって、より確からしい方の文字列の認識結果の重みを大きくすることになり、有効な照合結果の統合を行うことができる。なお、本発明において、「文字列」とは例えば広告主を表す漢字などの文字の列で示される「文字の列」であり、いわゆるイメージ・データの列である。また「文字コード列」とは例えば当該広告主を表す各文字を2値コードで表現する形で示した「コードの列」であり、いわゆるコード・データの列である。

【0009】

【発明の実施の形態】〔実施例1〕図1は本発明の一実施例構成を示す。

【0010】実施例の文字認識装置は、単語テーブル1、画像入力部2、文字列切り出し部3、文字切り出し部4、文字認識部5、広告主文字列信頼度計算部6、電話番号文字列信頼度計算部7、重み計算部8、広告主単語照合部9、電話番号単語照合部10、重み付け統合部11とで構成されている。

【0011】単語テーブル1には、広告主名と電話番号が対になって格納されている。

【0012】画像入力部2では、広告主と電話番号とが書かれている看板の画像（看板画像）の入力を行う。文字列切り出し部3では、画像入力部2で入力された看板画像から、区切れのないひとまとまりの行を文字列画像として切り出す。ここで使用する方法としては、例えば、書式に依存しない文字の形状及び配置に関する制約、および文字列の属性（文字サイズ、文字間隔等）に関する制約を与え、マルチエージェントシステムを利用して文字列を抽出する手法が提案されている（“マルチエージェントシステムを利用した文字領域抽出法”、隅谷倫子、行天啓二他、1995年電子情報通信学会総合大会、D-574,pp300）。

【0013】文字切り出し部4では、文字列切り出し部3で切り出された文字列画像全てについて、一文字ずつ文字画像を切り出し、文字サイズの算出を行う。文字認識部5では、全文字列について、切り出された各文字画像の認識を行い、認識結果文字コード列を出力する。文字切り出し部4および文字認識部5で使用する方法としては、例えば、文字線とその背景の濃度コントラストが高い、文字の外接矩形が正方形に近いものが多い、等の文字の普遍的な性質を利用することによって、雑音が多く条件の変化の激しい情景画像から文字を抽出、認識する方法が提案されている（“情景画像からの文字パターン抽出と認識”、大谷 淳、塩 昭夫、電子情報通信学会論文誌D Vol. J71-D No.6 pp.1037-1047）。

【0014】広告主文字列信頼度計算部6では、全文字列について、文字切り出し部4から出力された各文字の文字サイズのばらつきを用いて、広告主文字列としての認識結果の信頼度 $R_a$ を求める。具体的には、文字列中

の全文字の文字サイズ(=MAX「文字幅、文字サイズ」)の不偏標準偏差を平均で割った変動係数を算出する。変動係数は0~1の値を取り、1に近いほどばらつきが大きい。信頼度としては、変動係数を1からひいたものを用いる。

【0015】電話番号文字列信頼度計算部7では、全文字列について、電話番号文字列としての認識結果の信頼度 $R_t$ を求める。具体的には、認識結果に含まれる数字の割合(認識結果に含まれる数字の数を照合に使用する文字数で割った値)を信頼度とする。

【0016】重み計算部8では、広告主文字列信頼度計算部6と電話番号文字列信頼度計算部7とで求めた、広告主文字列候補の信頼度 $R_a$ と電話番号文字列候補の信頼度 $R_t$ とから、それぞれの照合結果に付ける重み $W_a$ 、 $W_t$ を、和が「1」になるように次式によって求める。

【0017】

$$W_a = R_a / (R_t + R_a) \dots \dots \dots (1)$$

$$W_t = R_t / (R_t + R_a) \dots \dots \dots (2)$$

広告主単語照合部9では、全文字列について、各文字列の認識結果文字コード列と単語テーブルの広告主文字コード列との単語照合を行い、電話番号単語照合部10では、全文字列について、各文字列の認識結果文字コード列と単語テーブルの電話番号文字コード列との単語照合を行い、それぞれ一致文字数を出力する。単語照合の方法としては、例えば、連想処理と統合処理の組み合わせにより、誤字・脱字・余剰時を許容できる連想統合型照合法が提案されている(“連想統合型照合による単語あいまい検索法”、松尾、佐藤、津田、第34回情処全大、4E-7, pp.1845-1846, 1987)。

【0018】重み付け統合部11では、広告主単語照合部9および電話番号単語照合部10より出力された広告主、電話番号それぞれの照合結果の一致文字数 $N_a$ 、 $N_t$ に、重み計算部8で計算された広告主、電話番号の重み $W_a$ 、 $W_t$ を付けて、統合結果のスコア $S$ を次式で求める。

【0019】

$$S = N_a \times W_a + N_t \times W_t \dots \dots \dots (3)$$

$S$ の値が最も高いものを、広告主候補として出力する。

【0020】このように構成した文字認識装置の動作および作用を説明する。図2ないし図4は一緒になって1つの図を構成し、図1に示した文字認識装置の動作を示すフローチャートである。

【0021】まず、ステップ20において看板画像を読み込む。

【0022】次にステップ21では、看板内の全文字領域について、1区切りのまとまった行を、文字列画像として切り出し、ステップ22では、文字列画像から一文字ずつ文字画像を切り出す。ステップ23で、ステップ22で切り出された文字画像それぞれの文字サイズを

算出し、ステップ24で文字認識を行い、認識結果文字コード列を生成する。ステップ25では、文字列として切り出した画像中の全ての文字について処理が終わったかどうかの判定を行い、ステップ26では、看板中の全ての文字列について処理が終わったかどうかを判定する。

【0023】ステップ27では、照合結果を統合したスコアの最大値 $S_{max}$ を0に設定しておく。

【0024】ステップ28で、一つ目の文字列を広告主文字列候補、ステップ29で二つ目の文字列を電話番号文字列候補とする。

【0025】ステップ30では、広告主文字列候補とした第 $m$ 文字列の全文字のサイズを読み込み、ステップ31で第 $m$ 文字列の認識結果文字コード列を読み込む。ステップ32では、電話番号文字列候補とした第 $n$ 文字列の全文字の認識結果文字コード列を読み込む。

【0026】次に、ステップ33で、広告主文字列候補の信頼度 $S_a$ として、第 $m$ 文字列の文字サイズの変動計数を求める。変動計数とは、標準偏差を平均で割った値で、ここでは全文字サイズの標準偏差と平均から求める。ステップ34では、電話番号文字列候補の信頼度 $S_t$ として、第 $n$ 文字列の認識結果文字コード列中に含まれる数字の割合を求める。具体的には、認識結果文字コード列中の数字の数を求め、それを文字数で割った値を $S_t$ とする。ステップ33とステップ34で求めた、広告主と電話番号それぞれの信頼度から、式(1)、式(2)により、それぞれの認識結果に掛ける重みをステップ35、ステップ36で算出する。

【0027】次に、ステップ37では、広告主 $m$ と電話番号 $n$ の組み合わせでの最大スコア $S(m, n)_{max}$ を0に初期化する。ステップ38とステップ39で、単語テーブル内の1対の広告主文字コード列と電話番号文字コード列について、それぞれ広告主文字列候補である第 $m$ 文字列の認識結果文字コード列と単語テーブル内の広告主文字コード列との一致文字数 $N_a$ 、電話番号文字列候補である第 $n$ 文字列の認識結果文字コード列と単語テーブル内の電話番号文字コード列との一致文字数 $N_t$ を計数する。

【0028】次にステップ40で、第 $m$ 文字列と第 $n$ 文字列の重み $W_a$ と $W_t$ 、広告主の一致文字数 $N_a$ と電話番号の一致文字数 $N_t$ から、式(3)によって統合した結果のスコア $S$ を算出し、ステップ41でそれが $S(m, n)_{max}$ より大きければ、ステップ42で $S(m, n)_{max}$ と入れ替え、ステップ43で広告主候補 $Adv(m, n)$ として格納する。ステップ44で、単語テーブル内の全ての広告主と電話番号の対についての照合が完了していれば、ステップ45にすすみ、完了していなければ、次の文字列対について、ステップ38からを繰り返す。

【0029】ステップ45では、広告主 $m$ と電話番号 $n$ の組み合わせについて、今までの文字列の組み合わせよ

り高いスコアが得られたかどうかを判定する。今までより高いスコアが得られていたら、ステップ46で最大スコアを書き換え、ステップ47で広告主候補を書き換える。

【0030】ステップ48で全ての文字列画像を電話番号候補として照合したかどうかの判定を行い、まだであれば、ステップ49～ステップ51で次の文字列を電話番号文字列候補とする。その際、広告主文字列と同じ行を除く。

【0031】ステップ52で全ての文字列画像を広告主候補として照合したかどうかの判定を行い、まだであればステップ53で次の文字列を広告主文字列候補とする。全ての文字列について処理が終わっていれば、ステップ54でAdvを広告主候補として出力し、処理を終了する。

【0032】〔実施例2〕図5は本発明の他の一実施例構成を示す。

【0033】実施例の文字認識装置は、単語テーブル1、画像入力部2、文字列切り出し部3、文字切り出し部4、文字認識部5、広告主文字列信頼度計算部6、電話番号文字列信頼度計算部7、重み計算部8、広告主単語照合部9、電話番号単語照合部10、重み付け統合部11、広告主文字列候補抽出部12、電話番号文字列候補抽出部13とで構成されている。

【0034】画像入力部2では、広告主と電話番号とが書かれている看板画像の入力を行う。文字列切り出し部3では、画像入力部2で入力された看板画像から、区切れないひとまとまりの行を文字列画像として切り出す。ここで使用する方法としては、例えば、書式に依存しない文字の形状及び配置に関する制約、および文字列の属性（文字サイズ、文字間隔等）に関する制約を与え、マルチエージェントシステムを利用して文字列を抽出する手法が提案されている（“マルチエージェントシステムを利用した文字領域抽出法”、隅谷倫子、行天啓二他、1995年電子情報通信学会総合大会、D-574, pp300）。

【0035】文字切り出し部4では、文字列切り出し部3で切り出された文字列画像全てについて、一文字分ずつ文字画像を切り出し、文字サイズの算出を行う。文字認識部5では、全文字列について、切り出された各文字画像の認識を行い、認識結果文字コード列を出力する。文字切り出し部4および文字認識部5で使用方法としては、例えば、文字線とその背景の濃度コントラストが高い、文字の外接矩形が正方形に近いものが多い、等の文字の普遍的な性質を利用することによって、雑音が多く条件の変化の激しい情景画像から文字を抽出、認識する方法が提案されている（“情景画像からの文字パターン抽出と認識”、大谷 淳、塩 昭夫、電子情報通信学会論文誌D Vol. J71-D No.6 pp.1037-1047）。

【0036】広告主文字列信頼度計算部6では、全文字

列について、文字切り出し部4から出力された各文字の文字サイズのばらつきを用いて、広告主文字列としての認識結果の信頼度Raを求める。具体的には、文字列中の全文字の文字サイズ（＝MAX「文字幅、文字サイズ」）の不偏標準偏差を平均で割った変動係数を算出する。変動係数は0～1の値を取り、1に近いほどばらつきが大きい。信頼度としては、変動係数を1からひいたものを用いる。

【0037】電話番号文字列信頼度計算部7では、全文字列について、電話番号文字列としての認識結果の信頼度Rtを求める。具体的には、認識結果に含まれる数字の割合（認識結果に含まれる数字の数を照合に使用する文字数で割った値）を信頼度とする。

【0038】重み計算部8では、広告主文字列信頼度計算部6と電話番号文字列信頼度計算部7とで求めた、広告主文字列候補の信頼度Raと電話番号文字列候補の信頼度Rtとから、それぞれの照合結果に付ける重みWa、Wtを、和が1になるように次式によって正規化を行う。

$$\begin{aligned} W_a &= R_a / (R_t + R_a) \dots \dots \dots (1) \\ W_t &= R_t / (R_t + R_a) \dots \dots \dots (2) \end{aligned}$$

広告主単語照合部9では、全文字列について、各文字列の認識結果文字コード列と単語テーブルの広告主文字コード列との単語照合を行い、電話番号単語照合部10では、全文字列について、各文字列の認識結果文字コード列と単語テーブルの電話番号文字コード列との単語照合を行い、それぞれ一致文字数を出力する。単語照合の方法としては、例えば、連想処理と統合処理の組み合わせにより、誤字・脱字・余剰時を許容できる連想統合型照合法が提案されている（“連想統合型照合による単語あいまい検索法”、松尾、佐藤、津田、第34回情処全大、4E-7, pp.1845-1846, 1987）。

【0040】重み付け統合部11では、広告主単語照合部9および電話番号単語照合部10より出力された広告主、電話番号それぞれの照合結果の一致文字数Na、Ntに、重み計算部8で計算された広告主、電話番号の重みWa、Wtを付けて、統合結果のスコアSを次式で求める。

$$S = N_a \times W_a + N_t \times W_t \dots \dots \dots (3)$$

Sの値が最も高いものを、広告主候補として出力する。

【0042】広告主文字列候補抽出部12では、各文字列のサイズを求め、サイズがもっとも大きい文字を含む文字列を広告主文字列候補として抽出する。文字列のサイズは、各文字列の文字画像のサイズの最大値とする。これは、看板に含まれる文字列の中で、広告主の文字列のサイズが最大であることが多いためである。

【0043】電話番号文字列候補抽出部13では、各文字列の認識結果の中で最も数字が多く含まれる文字列を



電話番号文字列候補として抽出する。これは看板に含まれる文字列の中で数字を最も多く含む文字列が電話番号であることが多いためである。

【0044】このように構成した文字認識装置の動作および作用を説明する。図6ないし図8は図5に示した文字認識装置の動作を示すフローチャートである。

【0045】まず、ステップ20において看板画像を読み込む。

【0046】次にステップ21では、看板内の全文字領域について、1区切りのまとまった行を、文字列画像として切り出し、ステップ22では、文字列画像から一文字分ずつ文字画像を切り出す。ステップ23で、ステップ22で切り出された文字画像それぞれの文字サイズを算出し、ステップ24で文字認識を行い、認識結果文字コード列を生成する。ステップ25では、文字列として切り出した画像中の全ての文字について処理が終わったかどうかの判定を行い、ステップ26では、看板中の全ての文字列について処理が終わったかどうかを判定する。

【0047】ステップ27では、照合結果を統合したスコアの最大値 $S_{max}$ を0に設定しておく。

【0048】ステップ28で、広告主文字列候補を抽出する。広告主文字列候補は全文字列の中でサイズが最も大きい文字を含む文字列とする。

【0049】ステップ29で電話番号文字列候補を抽出する。全文字列の認識結果に含まれる数字の数が最も多いものを電話番号文字列候補とする。

【0050】ステップ30では、広告主文字列候補の第 $m$ 文字列の全文字のサイズを読み込み、ステップ31で第 $m$ 文字列の認識結果文字コード列を読み込む。ステップ32では、電話番号文字列候補の第 $n$ 文字列の全文字の認識結果文字コード列を読み込む。

【0051】次に、ステップ33で、広告主文字列候補の信頼度 $S_a$ として、第 $m$ 文字列の文字サイズの変動計数を求める。変動計数とは、標準偏差を平均で割った値で、ここでは全文字サイズの標準偏差と平均から求める。ステップ34では、電話番号文字列候補の信頼度 $S_t$ として、第 $n$ 文字列の認識結果文字コード列中に含まれる数字の割合を求める。具体的には、認識結果文字コード列中の数字の数を求め、それを文字数で割った値を $S_t$ とする。ステップ33とステップ34で求めた、広告主と電話番号それぞれの信頼度から、式(1)、式(2)により、それぞれの認識結果に掛ける重みをステップ35、ステップ36で算出する。

【0052】次に、ステップ37では、広告主 $m$ と電話番号 $n$ の組み合わせでの最大スコア $S(m,n)_{max}$ を0に初期化する。ステップ38とステップ39で、単語テ

ル内の1対の広告主文字コード列と電話番号文字コード列について、それぞれ広告主文字列候補である第 $m$ 文字列の認識結果文字コード列と単語テーブル内の広告主文字コード列との一致文字数 $N_a$ 、電話番号文字列候補である第 $n$ 文字列の認識結果文字コード列と単語テーブル内の電話番号文字コード列との一致文字数 $N_t$ を計数する。

【0053】次にステップ40で、第 $m$ 文字列と第 $n$ 文字列の重み $W_a$ と $W_t$ 、広告主の一致文字数 $N_a$ と電話番号の一致文字数 $N_t$ から、式(3)によって統合した結果のスコア $S$ を算出し、ステップ41でそれが $S(m,n)_{max}$ より大きければ、ステップ42で $S(m,n)_{max}$ と入れ替え、ステップ43で広告主候補 $Adv(m,n)$ として格納する。ステップ44で、単語テーブル内の全ての広告主と電話番号の対についての照合が完了していれば、ステップ45にすすみ、完了していなければ、次の文字列対について、ステップ38からを繰り返す。

【0054】ステップ45では、広告主 $m$ と電話番号 $n$ の組み合わせについて、今までの文字列の組み合わせより高いスコアが得られたかどうかを判定する。今までより高いスコアが得られていたら、ステップ46で最大スコアを書き換え、ステップ47で広告主候補を書き換える。

【0055】ステップ48で $Adv$ を広告主候補として出力し、処理を終了する。

【0056】発明者らは、本発明の効果を確かめるために実験を行い、次の結果を得た。

【0057】実験には、デジタルビデオカメラで撮影した駅看板静止画像35枚を使用した。駅看板画像からの文字列切り出しと2値化とは、市販画像処理ソフトを用いて目視により行った。また、文字認識は市販文字認識ソフトを使用した。看板とタウンページ・データベースとの広告主には一部表記の不一致があるが、今回は照合方法の評価のため、タウンページ・データベース側の表記を看板に一致させて実験を行った。

【0058】文字列抽出に関しては、広告主文字列を抽出できたものが80%、電話番号文字列を抽出できたものが83%であった。

【0059】表1に広告主特定の実験結果を示す。広告主文字列の単文字認識精度は約40%であった。そのため広告主の単語照合のみを行っても特定精度は29%に過ぎない。一方、重みなしの統合による特定精度は69%であり、照合結果の統合により単語照合のみの場合に比べて特定精度が向上している。本発明の場合の特定精度は80%で、さらに高い精度が得られた。

【0060】

【表1】

広告主特定精度 (単位%)

照合 順位	重み付き統合 (本発明)	重みなし統合	広告主の 単語照合のみ
1 位	8 0	6 9	2 9
1 0 位以内	8 3	8 0	4 0

広告主文字列では切り出し誤り等に起因する誤認識の場合、認識結果に“.”や“-”が含まれるものが多く、タウンページ・データベース照合の際、それらの文字を含む広告主が上位になることが多い。また、電話番号文字列の認識結果に数字が少ない場合、数字として認識された文字にも誤認識が多く、正しい照合結果が得難い。本発明の場合では文字列認識結果の確からしさを重み付けにより反映させたため、誤認識文字の照合結果の重みが小さくなり、重みなしで統合した場合より高い特定精度が得られている。

【0061】上記において、本発明による文字認識装置と文字認識方法とについて説明したが、当該本発明による文字認識は、当該文字認識方法をデータ処理装置が実行可能にしたプログラムの形で保持することができる。このことから、本発明は当該プログラムを記録した記録媒体をも権利範囲に含むことは言うまでもない。

【0062】

【発明の効果】以上説明した如く、本発明によれば、看板などに対する文字認識精度をより向上することが可能となった。

【図面の簡単な説明】

【図1】本発明の一実施例構成を示す。

【図2】図1の構成に対応するフローチャートを示す。

【図3】図1の構成に対応するフローチャートを示す。

【図4】図1の構成に対応するフローチャートを示す。

【図5】本発明の他の一実施例構成を示す。

【図6】図5の構成に対応するフローチャートを示す。

【図7】図5の構成に対応するフローチャートを示す。

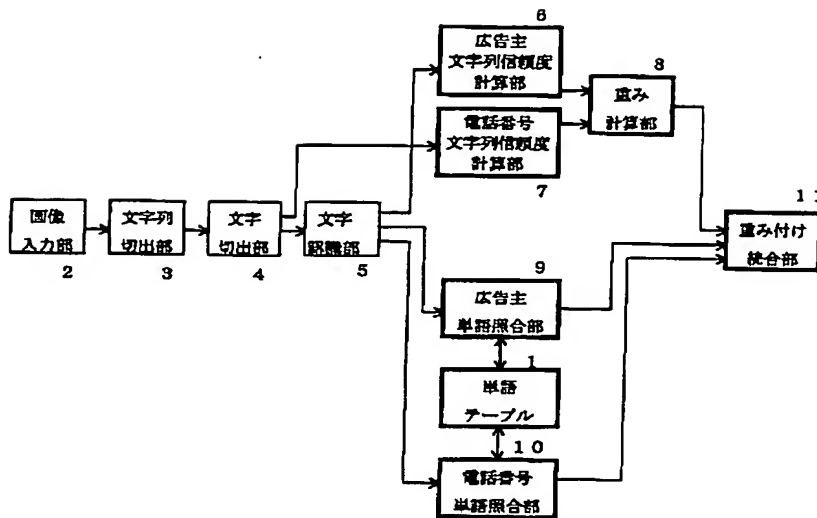
【図8】図5の構成に対応するフローチャートを示す。

【符号の説明】

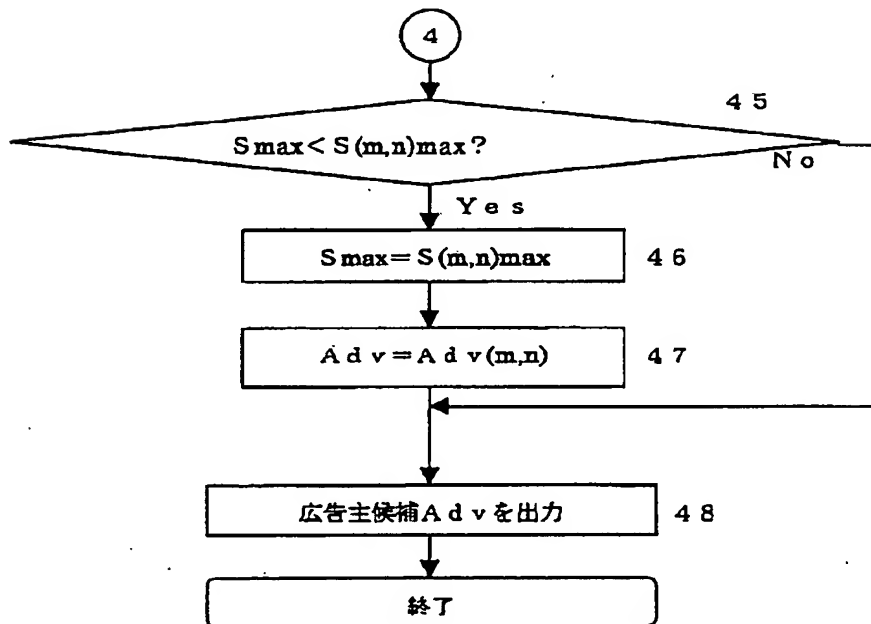
- 1 単語テーブル
- 2 画像入力部
- 3 文字列切り出し部
- 4 文字切り出し部
- 5 文字認識部
- 6 広告主文字列信頼度計算部
- 7 電話番号文字列信頼度計算部
- 8 重み計算部
- 9 広告主単語照合部
- 10 電話番号単語照合部
- 11 重み付け統合部
- 12 広告主文字列候補抽出部
- 13 電話番号文字列候補抽出部



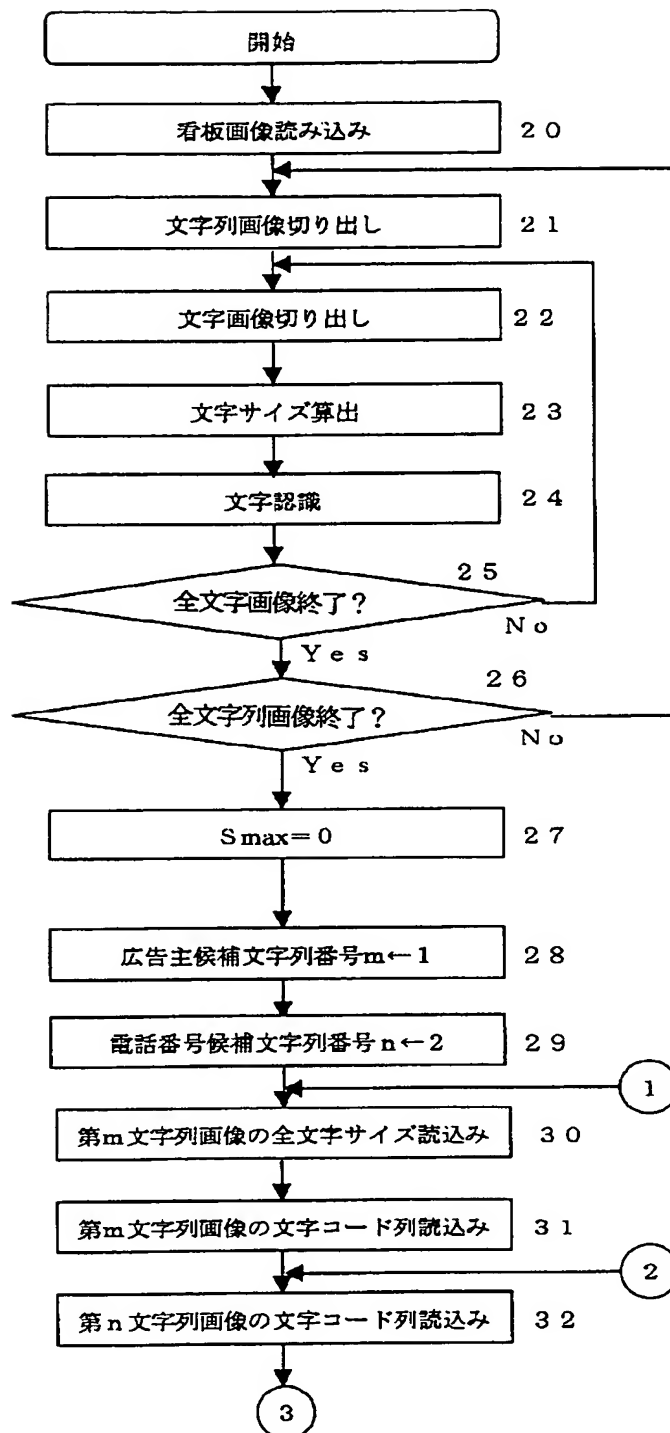
【図1】



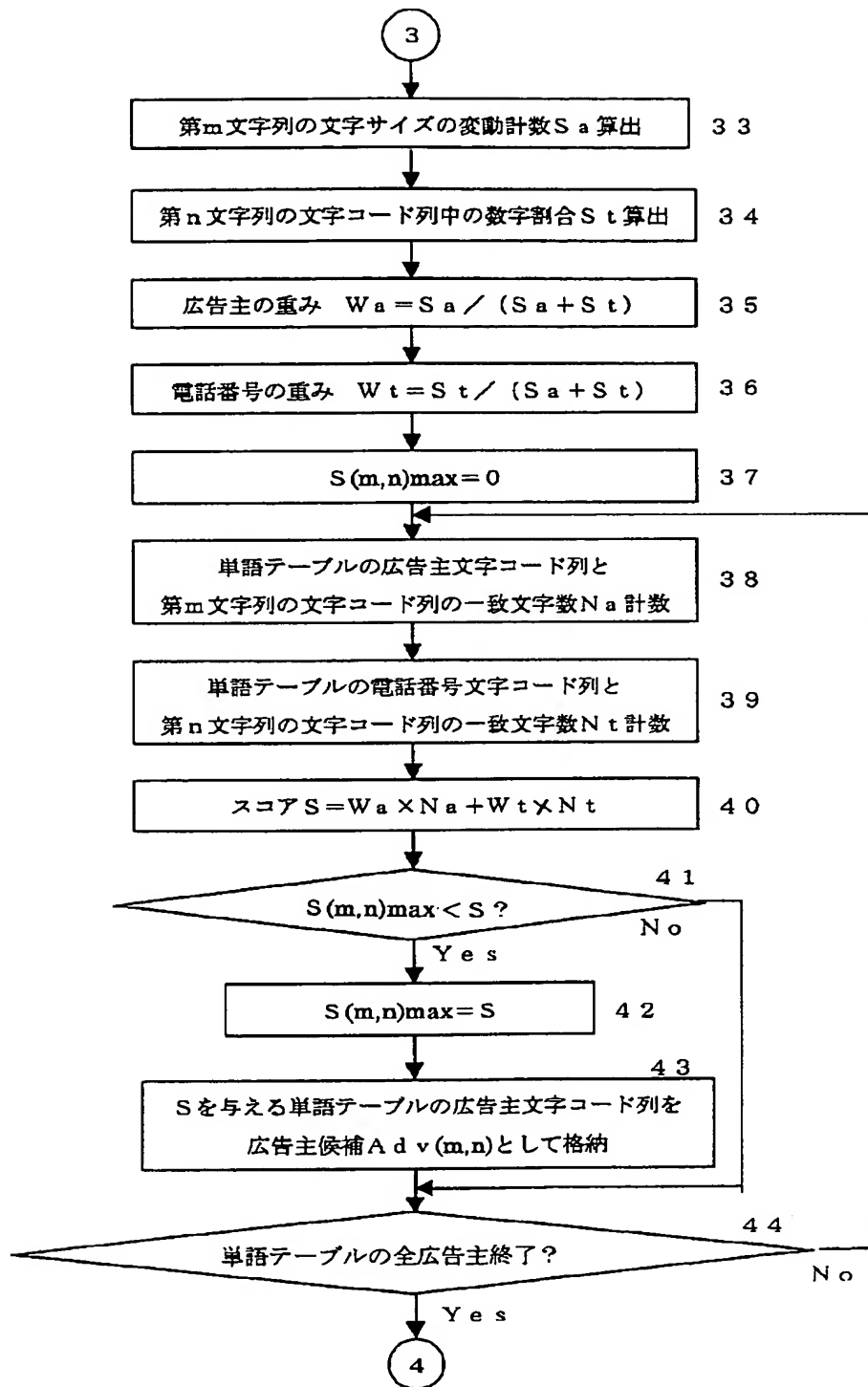
【図8】



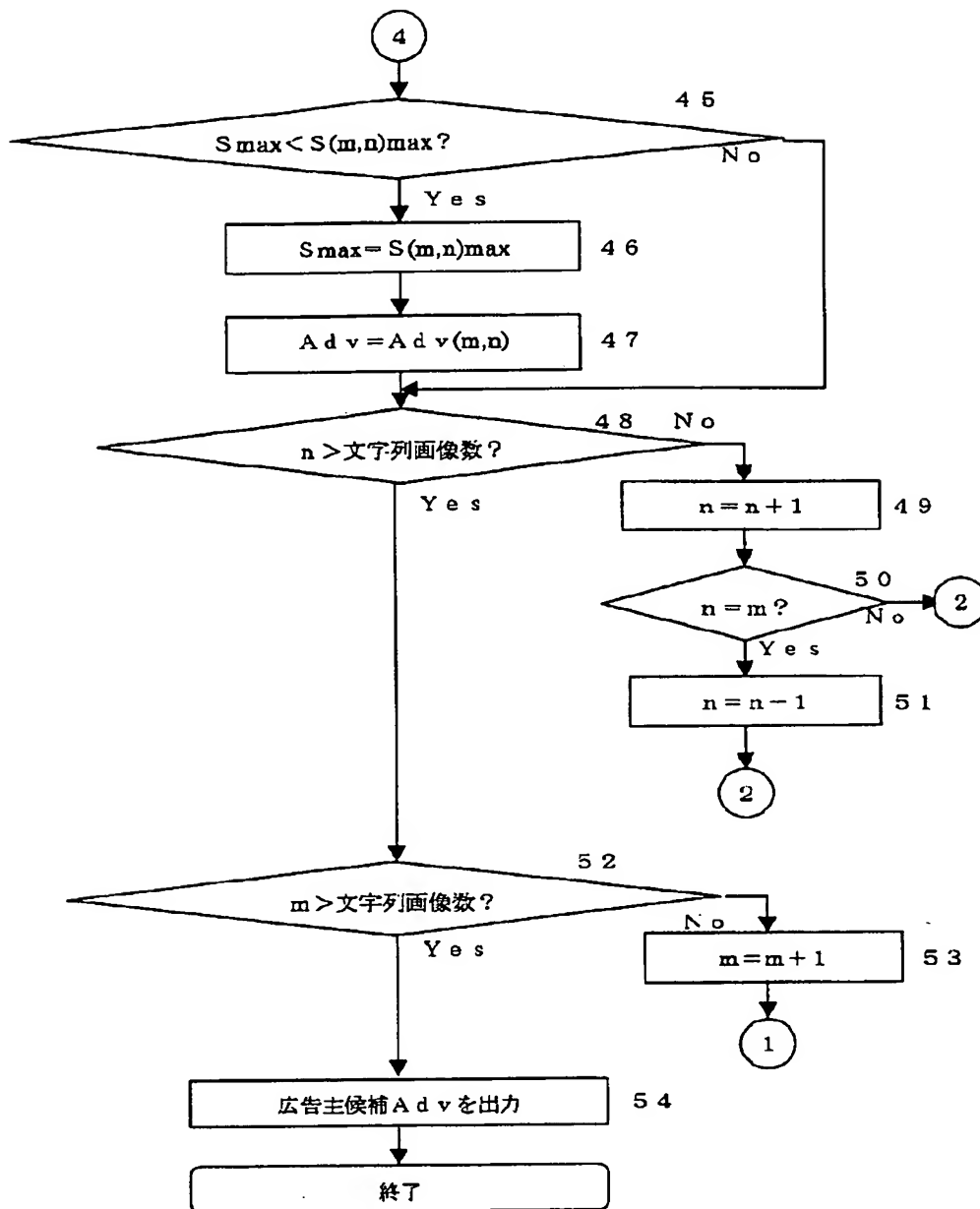
【図2】



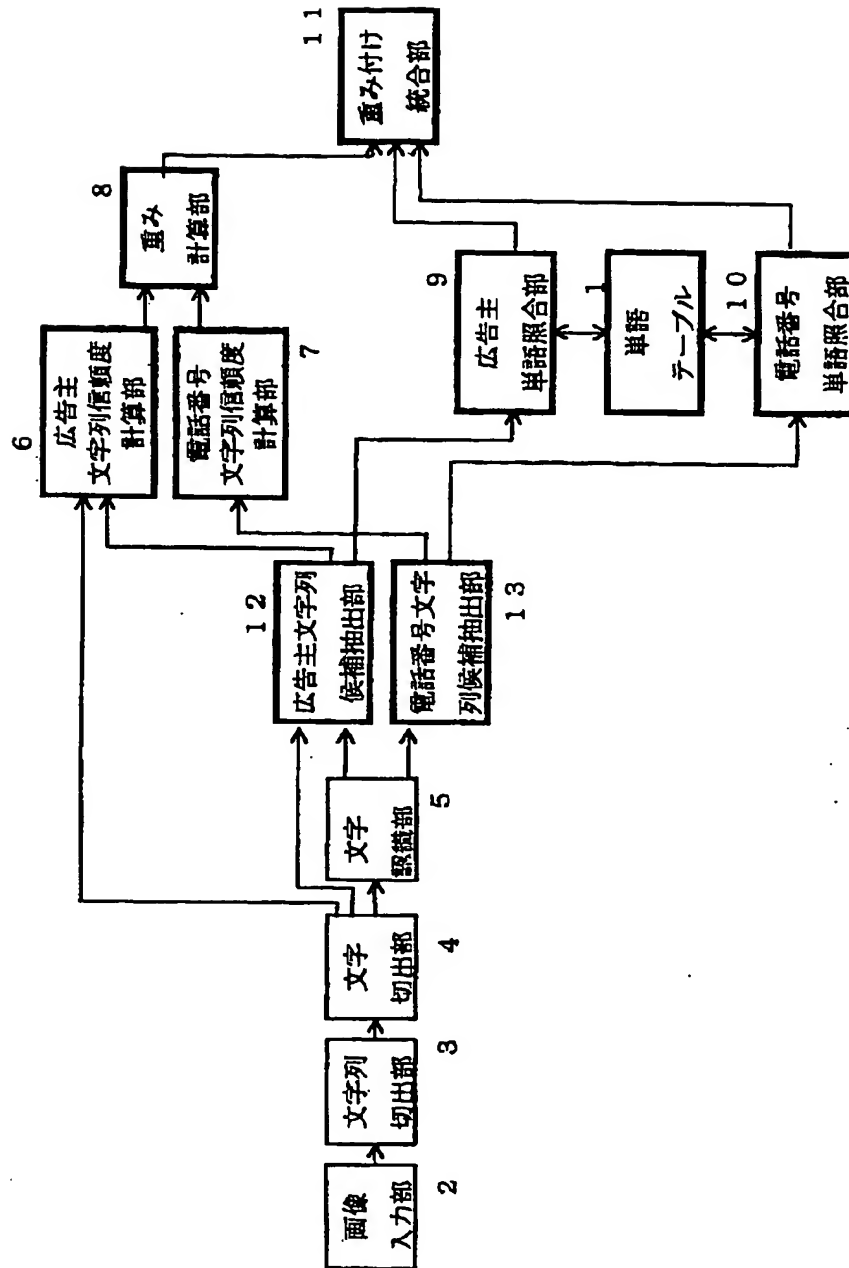
【図3】



【図4】

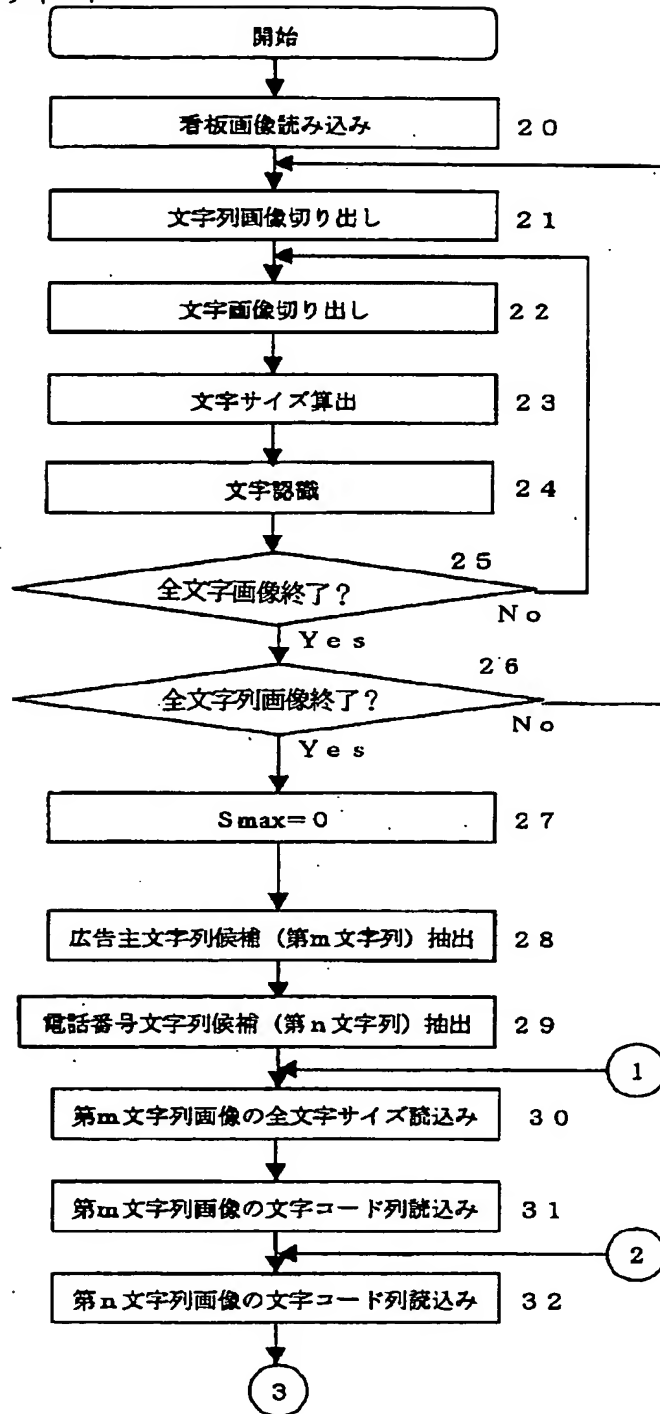


【図5】



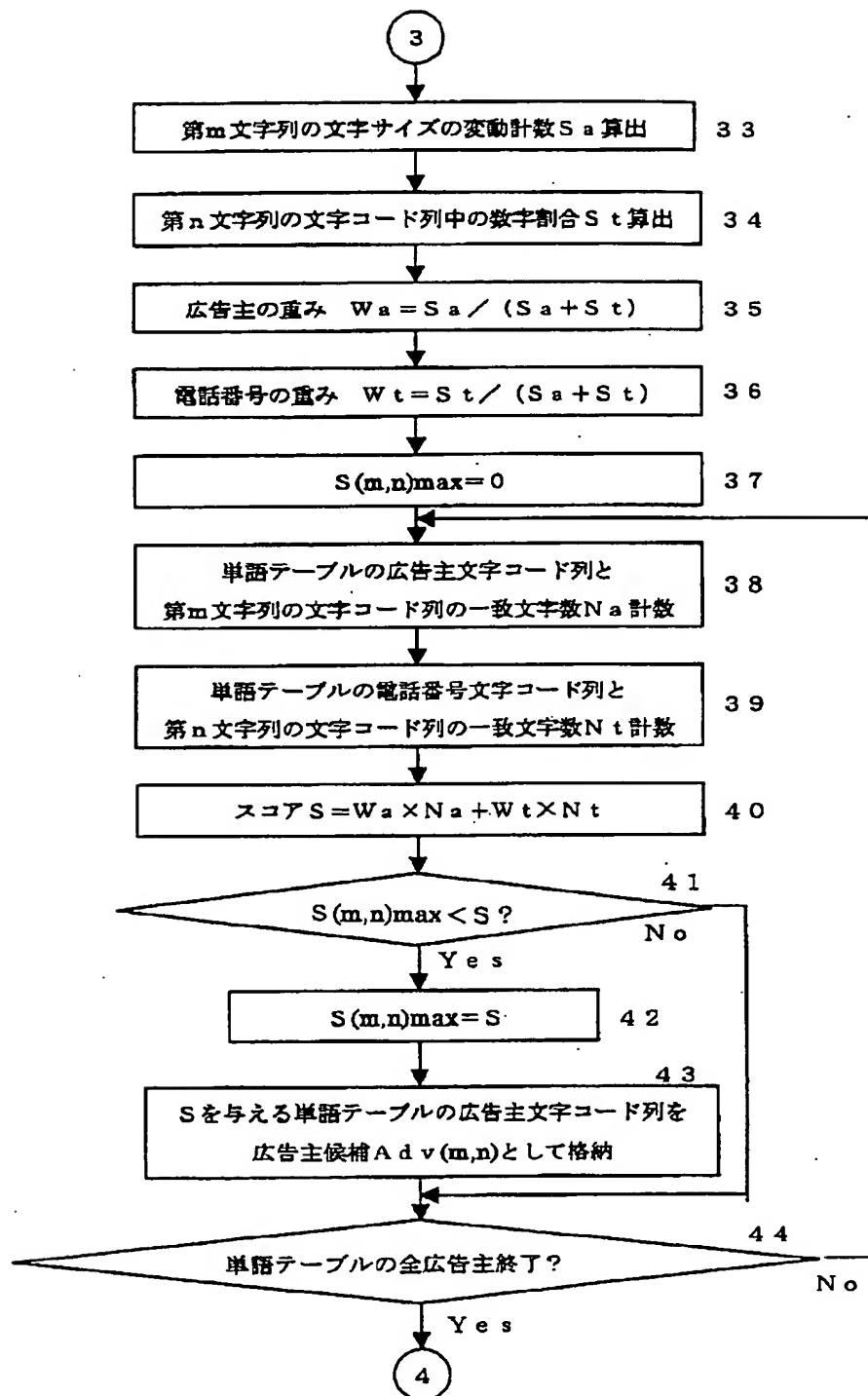
【図6】

—実施例2のフローチャート—





【図7】



フロントページの続き

(72)発明者 中村 修  
東京都新宿区西新宿三丁目19番2号 日  
本電信電話株式会社内

(72)発明者 杉村 利明  
東京都新宿区西新宿三丁目19番2号 日  
本電信電話株式会社内

(56)参考文献 特開 平11-143984 (J P, A)  
特開 平6-180767 (J P, A)  
特開 平4-328692 (J P, A)  
特開2000-48169 (J P, A)

(58)調査した分野(Int.Cl.<sup>7</sup>, D B名)  
G06K 9/00 - 9/82